

# Social Determinants of Health and High-Cost Utilization Among Commercially Insured Population

Elena Andreyeva, PhD; Winnie Chi, PhD; Yongkang Zhang, PhD; Rainu Kaushal, MD, MPH; and Kevin Haynes, PharmD, MSCE

The US health care system has been transitioning toward value-based care models that offer provider incentives to improve health outcomes at a lower cost.<sup>1-6</sup> These efforts drive an increased understanding that social factors, such as occupation and residential environment, are important drivers of health care utilization and outcomes.<sup>7-9</sup> Collectively, these nonmedical elements are called social determinants of health (SDOH)<sup>10,11</sup>: environmental conditions in which people live and work that may affect their health and other quality-of-life outcomes. SDOH encompass a range of conditions, including economic stability, education, health and health care services, social and community considerations, and neighborhood and physical environment.

Although high-cost utilizers are a relatively small proportion of the population, they consume a large proportion of health care resources. Previous studies have found that certain SDOH, including low income, lower educational attainment, food insecurity, and not owning a home, are major risk factors for being a high-cost utilizer.<sup>12</sup> Greater risks of morbidity and mortality,<sup>13-15</sup> coupled with nonadherence to treatment or delayed treatment due to economic instability, intensify the likelihood that individuals with disadvantaged SDOH will become high-cost utilizers.<sup>13,14,16,17</sup>

In high-cost populations, the majority of spending is due to utilization of institutional care<sup>18</sup>; however, literature shows that the types of health care services change with age.<sup>19</sup> Utilization of acute care services, such as inpatient hospitalizations and emergency department (ED) visits, increases with age, whereas the use of primary care services decreases with age.<sup>19</sup> This difference also results in part from underlying health status, with many older adults, but few children, having chronic conditions.<sup>19</sup> Although lower socioeconomic status (SES) generally exposes individuals to a greater variety of risk factors, its impact on the risk of becoming a high-cost utilizer may differ by age groups, given different utilization patterns and underlying health conditions.<sup>20</sup> An important question is whether the relationship between SDOH and high-cost status varies across age, but this has not been widely studied in the literature, especially within the more homogenous commercially insured population.

## ABSTRACT

**OBJECTIVES:** To assess the impact of adding neighborhood social determinants of health (SDOH) data to demographic and clinical characteristics for predicting high-cost utilizers and to examine variations across age groups.

**STUDY DESIGN:** Using US Census data and 2017-2018 commercial claims from a large national insurer, we estimated association between neighborhood-level SDOH and the probability of being a high-cost utilizer.

**METHODS:** Observational study using administrative claims from a national insurer and US Census data. Data included a 50% random sample of commercially insured individuals who were younger than 89 years and had 1 year of continuous eligibility in 2017 and at least 30 days in 2018. Probit models assessed impact of SDOH and neighborhood conditions on predicting cost status.

**RESULTS:** SDOH did not improve predictive power of evaluated models. However, disadvantaged neighborhood residence was still associated with being a high-cost utilizer. Adults 65 years and older in disadvantaged neighborhoods had increased likelihood of high-cost utilization. Children and younger adults in disadvantaged neighborhoods had lower risk of becoming high-cost utilizers.

**CONCLUSIONS:** Policy makers and industry stakeholders should be aware of the mechanisms behind the relationship between neighborhood social conditions and health outcomes and how the relationship differs across age groups.

*Am J Manag Care.* 2023;29(7):e199-e207. doi:10.37765/ajmc.2023.89399

## TAKEAWAY POINTS

As enrollment in Medicare Advantage plans continues to increase, it is crucial for stakeholders to recognize the importance of social determinants of health in older individuals' health status and subsequent health care utilization and costs. This study found that residence in more disadvantaged neighborhoods is associated with a higher likelihood of becoming a high-cost utilizer among older commercially insured adults and Medicare Advantage beneficiaries.

- ▶ Medicare Advantage plans may offer supplemental services to address unmet social needs, in addition to medical care, especially for members residing in more disadvantaged neighborhoods.
- ▶ Our findings support the growing trend among health insurers to utilize a more holistic approach to assessing individuals' well-being.

In this study, we aim to fill the gap in literature by evaluating how the relationship between SDOH and the risk of becoming a high-cost utilizer differs across age groups. Specifically, we aim to (1) develop a predictive model to assess whether adding neighborhood SDOH to demographic and clinical characteristics can improve the accuracy of predicting high-cost utilizers among commercially insured members in an overall population and by age groups, and (2) determine whether the relationship between SDOH and the risk of becoming a high-cost utilizer varies by age group.

Our evaluation demonstrates that SDOH could provide additional valuable information to predict the risk of becoming a high-cost utilizer, especially in situations in which claims history is lacking or incomplete, such as for new individual enrollees or a new employer group. In addition, better recognition of relevant neighborhood social conditions has the potential to enhance the ability of providers, payers, and policy makers to target social services and interventions to reduce the risk of becoming a high-cost utilizer.

## METHODS

### Sample and Data

The study used a sample of commercially insured members from a large national health insurer with approximately 40 million members. Members were included if they were younger than 89 years with a full year of continuous enrollment in 2017 and at least 30 days of continuous enrollment in 2018. Requiring a full year of enrollment in 2017 allowed us to get a complete picture of members' health characteristics in the base year while not biasing against including members who might have died or lost their health insurance in the follow-up year. To optimize the computation time, we randomly selected 50% of the population who met these inclusion criteria, bringing the sample size to approximately 3.1 million members including Medicare Advantage members. Two data sources were used: commercial medical and pharmacy administrative claims from the health insurer and the 2017 American Community Survey (ACS) 5-year estimates at the US Census block group level (2013-2017). We merged the 2 data sets using each member's 9-digit zip code of residence matched to a Census block group using a ZIP+4 crosswalk database.

### Variables

The dependent variable of interest was an indicator variable for incurring per-member per-month (PMPM) health care expenditures in 2018 in the top decile (90th percentile) in the member's state of residence to account for wide geographic variations in health care costs across the United States. For instance, if an individual's state of residence was New York based on health plan enrollment data, they would be assigned an indicator variable equal to 1 if their 2018 total PMPM cost was greater than the 90th percentile of the 2018 total PMPM cost

of New York; otherwise, it would be 0. We applied this method for all individuals residing in states with at least 10,000 members in our full sample. For all individuals residing in states with fewer than 10,000 members in our study sample (< 0.3% of our total sample), we assigned a 0 or 1 value based on being below or above the 90th percentile of the 2018 total national PMPM cost (\$1128.89) to account for the fact that the cost distribution from our data might not be representative for the states with fewer observations. Cancer- and transplant-related costs were excluded due to extremely high cost; otherwise, our top decile would be overrepresented by members with cancer and transplants. For age subsample analyses, we used the 90th percentile for the corresponding age subgroup as the cutoff.

The main independent variable of interest is a composite SDOH variable measured by the Agency for Healthcare Research and Quality (AHRQ) SES index.<sup>21</sup> AHRQ developed the SES index specifically for use with administrative claims data that lack individual-level SES measures. The index is based on the member's Census block group of residence and includes the following ACS variables: unemployment rate, percentage of people living below poverty level, median household income, median value of owner-occupied dwellings, percentage of people 25 years and older with less than a 12th-grade education, percentage of people 25 years and older with at least 4 years of college, and percentage of households with at least 1 person per room (crowding). Higher index scores indicate higher SES. The SES index was constructed for all block groups using the 2017 ACS. Following a 2008 AHRQ report prepared by RTI International investigators, we stratified all block groups in the country into quartiles: 1 (lowest SES, reference group), 2, 3, and 4 (highest SES).<sup>21</sup>

Our analysis included a large set of individual- and neighborhood-level characteristics that may correlate with high-cost status as independent variables. Individual characteristics include 2017 continuous age measure, dummy for being female, utilization (counts of inpatient stays; outpatient, specialist, wellness, and routine visits; and prescription fills), costs (total costs associated with the aforementioned utilization), and comorbidities (modified Charlson Comorbidity Index score and indicators for having a number of comorbidities).<sup>22</sup> We used *International Classification of Diseases, Tenth Revision (ICD-10)* codes to identify comorbidities

in patients' claims. An individual was assigned a diagnosis if they had 1 inpatient or 2 outpatient visits with that *ICD-10* code in 2017.

### Statistical Analysis

We predicted high-cost status using the following 4 regressions by sequentially adding different sets of predictors:

$$y_{is} = \gamma \text{gender}_i + \delta \text{age}_i + \mu_s + \epsilon_{is}$$

$$y_{is} = \gamma \text{gender}_i + \delta \text{age}_i + \alpha \text{SEScategory}_b + \mu_s + \epsilon_{is}$$

$$y_{is} = \gamma \text{gender}_i + \delta \text{age}_i + \omega X_i + \mu_s + \epsilon_{is}$$

$$y_{is} = \gamma \text{gender}_i + \delta \text{age}_i + \omega X_i + \alpha \text{SEScategory}_b + \mu_s + \epsilon_{is}$$

where  $y$  is an indicator variable that equals 1 if the individual ( $i$ ) is a high-cost utilizer in 2018 and 0 otherwise.  $X$  is a vector of  $i$ 's 2017 costs, utilization patterns, and clinical characteristics. *SEScategory* is a set of dummy variables for whether an individual resides in a block group that belongs to SES quartiles 2, 3, or 4 (1 is the reference group);  $\mu$  is state fixed effects; and  $\epsilon$  is the error term.

We estimated the regressions using probit models and reported marginal effects in the analysis tables. We also reported heteroscedasticity-robust SEs. We then stratified our sample by age groups (< 18, 18-44, 45-64, and 65-89 years) to see whether the association between high-cost status and SES category varies by age. The analysis was performed using Stata 16 (StataCorp LLC).

### Model Validation

To compare the performance of different model specifications, we performed 10-fold cross validation. We split our data into 10 equal-sized parts and built 10 models; each model was trained on the first 9 parts and tested on the 10th part. In other words, we used the 9 training groups to estimate probit regressions 1, 2, 3, and 4 and then used the 10th testing group to predict the value of the dependent variable using the coefficients from the probit regressions on the training data. This exercise was repeated 10 times, once for each "testing" model. We then reported the mean of the area under the curve (AUC) of the receiver operating characteristic curve from the 10 exercises, which allowed us to visualize how much each of our models was capable of distinguishing between high-cost utilizers and those who were not. We also reported Brier score, which evaluates the accuracy of probabilistic predictions, with a lower value indicating a more accurate prediction.

## RESULTS

### Population Characteristics, Health Care Resource Utilization, and Costs

**Table 1** presents sociodemographic characteristics, health care resource utilization, and costs by 2018 high-cost status. The 2018 mean PMPM costs were \$4042 and \$181 for high-cost and non-high-cost utilizers, respectively. High-cost utilizers were more likely than non-high-cost utilizers to be female (57% vs 50%, respectively) and older (mean age, 49.4 vs 37.9 years). High-cost utilizers had at

least 3 times higher mean visits per month across settings in 2017 (inpatient stays, 0.26 vs 0.05; ED visits, 0.45 vs 0.15; and outpatient visits, 24.98 vs 7.38). Consequently, they had higher 2017 mean annual inpatient costs (\$6488 vs \$816), ED costs (\$976 vs \$277), outpatient costs (\$10,536 vs \$1533), and specialist costs (\$10,776 vs \$1774) than the non-high-cost utilizers.

### SES Distribution

The largest proportion of individuals in our sample resided in the least disadvantaged community (SES quartile [Q] 4, 38.8%;  $n = 1,202,197$ ), followed by Q3 (27.4%;  $n = 850,977$ ), Q2 (21.3%;  $n = 661,937$ ), and the most disadvantaged, Q1 (12.5%;  $n = 386,148$ ).

According to **Figure 1**, more high-cost utilizers resided in the most disadvantaged neighborhoods than non-high-cost utilizers (13.4% vs 12.3%), whereas fewer high-cost utilizers resided in the least disadvantaged neighborhoods than non-high-cost utilizers (36.9% vs 39.0%). All differences are significant at  $P < .001$ . Overall, both high-cost and non-high-cost individuals in our sample appear to be fairly equally distributed across the 4 types of neighborhoods.

### Model Performance With SES Factors

**Figure 2** presents the performance of different prediction models. Adding SES index did not improve the performance of the base model (age-gender model). The mean AUC was 0.666, and the Brier score was 0.0871 in both age-gender and age-gender-SDOH models. The SES index also did not change the performance of the claims history model, in which the mean AUC was 0.819 and the Brier score was 0.0710 with and without SDOH. Addition of the SES index also did not change the model performance in the age subgroup analysis.

### Marginal Effect of SES Factors on High-Cost Status

Although neighborhood-level SES did not change the model's performance, residence in less disadvantaged neighborhoods (Q3 and Q4) was associated with a lower risk of becoming a high-cost utilizer relative to residence in the most disadvantaged neighborhoods (Q1). As shown in **Table 2**, residing in Q3 and Q4 neighborhoods was associated with 0.502 and 0.677 percentage points lower probability, respectively, of being a high-cost utilizer in the age-gender model ( $P < .001$ ). Similarly, in the model with clinical characteristics, residing in Q3 and Q4 neighborhoods was associated with 0.328 and 0.577 percentage points lower probability of being a high-cost utilizer ( $P < .001$ ). However, residence in less disadvantaged neighborhoods had a smaller effect on the probability of being a high-cost utilizer than did utilization or clinical characteristics.

### Marginal Effect of SES Factors on High-Cost Status by Age Groups

Interestingly, we found that the relationship between residence in less disadvantaged neighborhoods and the risk of becoming a high-cost utilizer was directionally different between the subsamples of children and older adults. **Table 3** shows the marginal effects

**TABLE 1.** Summary Statistics by High-Cost Utilizer Status in 2017

Characteristic	Top 10% high-cost utilizers (n = 308,389)			Non-top 10% high-cost utilizers (n = 2,792,870)			P
	Mean	SD	Median	Mean	SD	Median	
2018 total PMPM cost	\$4042	\$7112	\$2238	\$181	\$246	\$79	<.001
<b>Variables in 2017</b>							
<b>Demographic characteristics</b>							
Age in years	49.36	18.22	52.00	37.91	20.92	39.00	<.001
Female, %	57.10	49.50		49.90	50.00		<.001
<b>Neighborhood ACS characteristics, %</b>							
SES index score	55.70	6.11	55.12	55.94	6.01	55.46	<.001
SES quartile 1	13.40	34.10		12.30	32.90		<.001
SES quartile 2	22.40	41.70		21.20	40.90		<.001
SES quartile 3	27.20	44.50		27.50	44.60		.006
SES quartile 4	36.90	48.20		39.00	48.80		<.001
Education < 9 years	3.85	5.28	2.15	3.83	5.47	2.06	.006
Education high school and above	90.00	9.20	92.50	90.20	9.30	92.90	<.001
Education high school only	26.60	13.30	26.60	26.00	13.20	25.80	<.001
Education GED	3.79	3.74	2.83	3.63	3.62	2.68	<.001
Employed in white collar occupations (n=308,362)	66.10	18.40	66.90	66.80	18.20	67.80	<.001
Unemployment rate	5.85	5.20	4.64	5.69	5.02	4.53	<.001
Population below FPL	11.40	10.90	7.99	10.90	10.60	7.58	<.001
Population with income < 150% of FPL	18.90	15.10	14.80	18.20	14.70	14.10	<.001
Population with income < 200% of FPL	26.80	18.00	23.00	25.90	17.70	22.00	<.001
Population in labor force	35.90	0.11	34.90	35.30	10.60	34.40	<.001
HHs with public assistance	2.08	3.27	0.96	2.00	3.16	0.88	<.001
HHs with SSI	4.71	5.10	3.31	4.48	4.87	3.16	<.001
Per capita income (n=308,385; n=2,792,847)	\$36,825	\$20,817	\$31,628	\$36,989	\$19,890	\$32,245	<.001
Median family income	\$86,735	\$40,739	\$76,652	\$87,985	\$40,112	\$78,645	<.001
High-need population	22.20	7.36	21.70	22.30	7.28	21.80	<.001
Income disparity index (n=308,248; n=2,792,023)	1.86	1.31	188.00	1.78	1.28	180.60	<.001
Single-parent HHs with kids (n=308,154; n=2,791,916)	29.20	19.10	25.60	28.30	18.80	24.80	<.001
Population without car	6.51	10.60	3.08	6.00	9.91	2.87	<.001
Population without complete plumbing	0.34	1.29	0.00	0.33	1.30	0.00	.004
Population without telephone (n=307,001; n=2,780,944)	1.98	2.96	1.11	1.93	2.91	1.11	<.001
Crowding	2.14	4.12	0.00	2.20	4.32	0.00	<.001
Mobile housing	5.29	10.70	0.00	4.95	10.20	0.00	<.001
Vacancy rate	9.57	10.20	7.09	9.13	9.89	7.09	<.001
Housing tenure	70.90	22.50	76.90	71.60	22.50	77.70	<.001
Population in group quarters	1.45	5.48	0.00	1.43	5.47	0.00	.035
Median home value	\$299,622	\$300,635	\$196,200	\$298,065	\$286,378	\$201,850	.004
Median gross rent (n=299,518; n=2,706,199)	\$1165	\$552	\$1006	\$1171	\$543.40	\$1021	<.001
Owner cost with mortgage (n=307,752; n=2,706,199)	\$1704	\$782.80	\$1456	\$1712	\$763.70	\$1485	<.001
Housing costs to mortgage (n=307,906; n=2,788,855)	22.40	4.60	21.50	22.30	4.56	21.40	<.001
Black	9.47	17.40	2.18	9.04	16.70	2.18	<.001
Hispanic	9.63	14.90	3.94	9.88	15.60	3.94	<.001
Asian	4.85	9.50	1.10	5.22	10.20	1.10	<.001
White non-Hispanic	73.50	25.00	81.90	73.30	25.30	81.90	<.001
HHs with English spoken at home less than very well	2.31	4.82	0.34	2.40	5.00	0.34	<.001

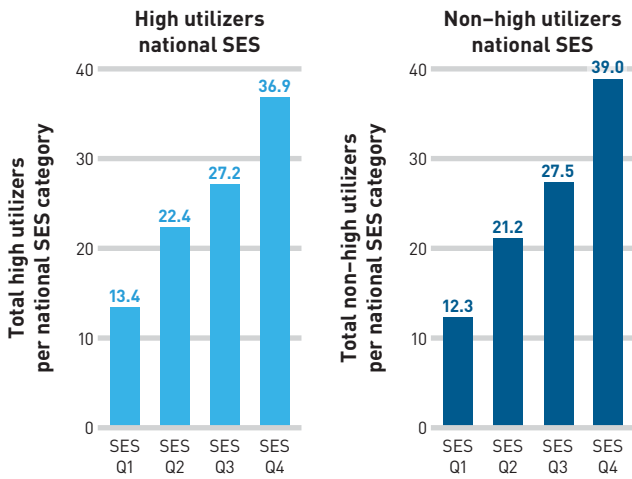
(continued)

**TABLE 1.** (Continued) Summary Statistics by High-Cost Utilizer Status in 2017

Characteristic	Top 10% high-cost utilizers (n = 308,389)			Non-top 10% high-cost utilizers (n = 2,792,870)			P
	Mean	SD	Median	Mean	SD	Median	
<b>Variables in 2017 (continued)</b>							
<b>Neighborhood ACS characteristics, % (continued)</b>							
Population aged ≥ 65 years	16.70	9.02	15.40	16.20	8.53	15.10	<.001
Population aged ≤ 17 years	21.90	7.61	22.10	22.30	7.60	22.50	<.001
Workers with long commute (n=308,387; n=2,792,852)	8.43	8.34	6.04	8.32	8.29	6.04	<.001
Workers with car (n=308,388; n=2,792,868)	87.80	14.00	91.50	88.00	13.40	91.50	<.001
<b>Clinical conditions (%)</b>							
Modified CCI score, mean	2.30	2.77	1.00	0.63	1.27	0.00	<.001
Alzheimer disease	0.40	6.34		0.13	3.54		<.001
Arthritis	15.80	36.50		3.97	19.50		<.001
Asthma	6.09	23.90		2.06	14.20		<.001
Congestive heart failure	4.50	20.70		0.60	7.75		<.001
Chronic obstructive pulmonary disease	5.89	23.60		1.00	9.97		<.001
Cancer	9.87	29.80		2.34	15.10		<.001
Cardiovascular	12.40	32.90		3.00	17.10		<.001
Crohn disease	1.32	11.40		0.10	3.09		<.001
Depression	12.30	32.90		3.93	19.40		<.001
Diabetes	20.70	40.50		4.79	21.40		<.001
Adrenal insufficiency	5.94	23.60		1.00	9.94		<.001
Epilepsy	1.53	12.30		0.28	5.30		<.001
Hepatitis C virus	0.78	8.79		0.14	3.77		<.001
HIV	1.11	10.50		0.02	1.49		<.001
Multiple sclerosis	1.30	11.30		0.06	2.45		<.001
Mental illness	0.69	8.28		0.12	3.50		<.001
Pregnancy	0.85	9.19		0.93	9.59		<.001
Schizophrenia	0.25	4.98		0.04	2.08		<.001
Sickle cell disease	0.07	2.59		0.01	1.17		<.001
Stroke	1.30	11.30		0.24	4.92		<.001
<b>2017 utilization (n)</b>							
ED visits	0.45	1.21	0.00	0.15	0.48	0.00	<.001
Inpatient stays	0.26	0.80	0.00	0.05	0.27	0.00	<.001
Office visits	24.98	34.73	7.00	7.38	11.99	3.00	<.001
Other outpatient visits	9.28	8.56	15.00	3.71	4.22	3.00	<.001
Wellness visits	0.03	0.18	0.00	0.02	0.13	0.00	<.001
Rx fills	21.18	20.16	16.00	6.79	10.20	2.00	<.001
Specialist visits	27.25	34.79	17.00	8.26	12.09	4.00	<.001
Routine physical visits	0.42	0.61	0.00	0.46	0.67	0.00	<.001
<b>2017 cost (\$)</b>							
Total specialist	\$10,776	\$35,172	\$2951	\$1774	\$5289	\$529	<.001
Total inpatient stay	\$6488	\$37,975	\$0	\$816	\$7902	\$0	<.001
Total ED	\$976	\$3287	\$0	\$277	\$1313	\$0	<.001
Total other outpatient	\$10,537	\$35,601	\$2399	\$1533	\$5273	\$264	<.001
Total office visit	\$1017	\$1068	\$738	\$385	\$450	\$259	<.001
Total wellness	\$4	\$22	\$0	\$2	\$14	\$0	<.001
Total Rx	\$8662	\$22,683	\$1437	\$665	\$2579	\$52	<.001

ACS, American Community Survey; CCI, Charlson Comorbidity Index; ED, emergency department; FPL, federal poverty level; GED, General Educational Development; HH, household; PMPM, per-member per-month; Rx, prescription; SES, socioeconomic status; SSI, Supplemental Security Income.

**FIGURE 1.** Percentage of Individuals Residing in Each SES Quartile by High-Cost Status



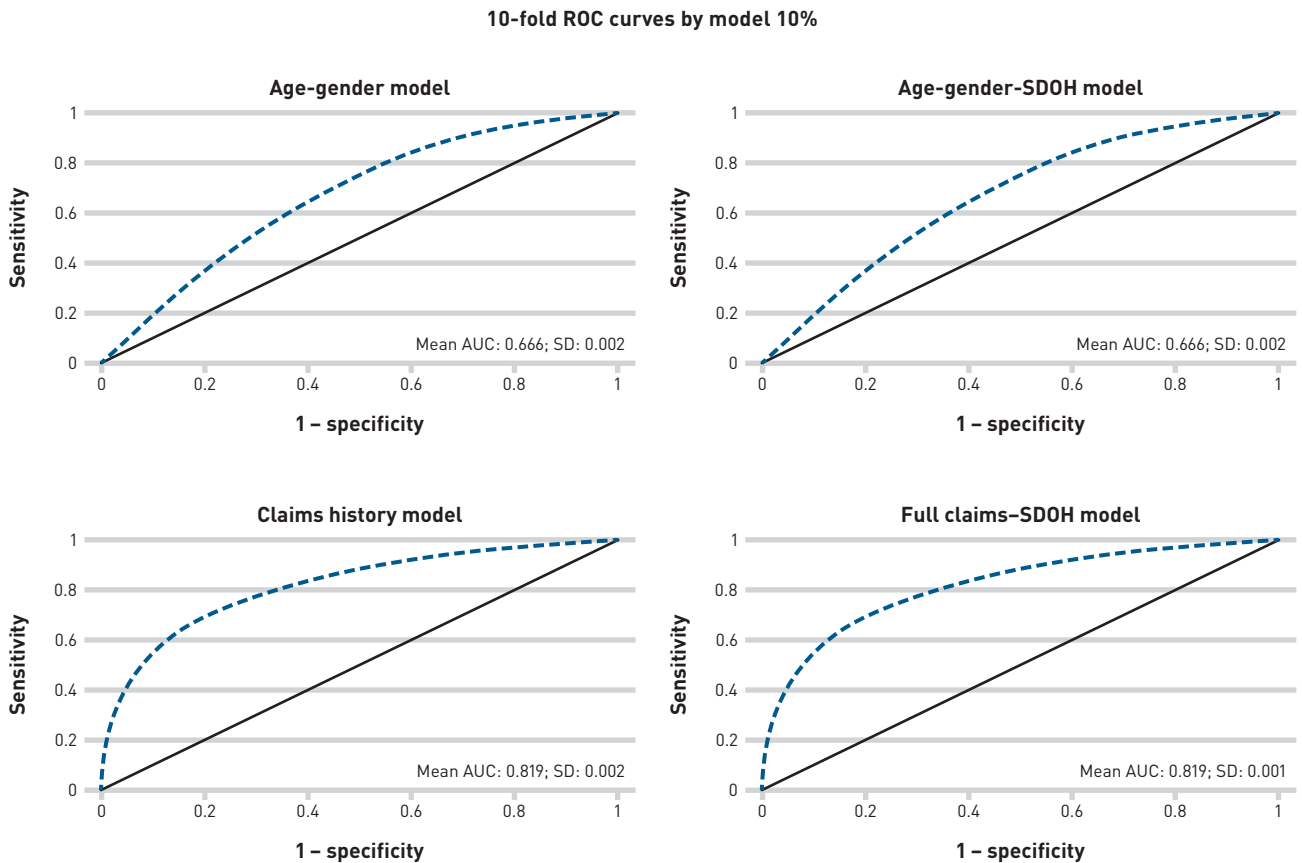
Q, quartile; SES, socioeconomic status.

from the age-gender-SDOH model and the claims history-SDOH model for each subsample. Overall, children residing in higher-SES neighborhoods were at a higher risk of becoming high-cost utilizers in the claims history-SDOH model. In the age-sex-SDOH model, residence in a Q2, Q3, or Q4 neighborhood was positively associated with becoming a high-cost utilizer. The results change significantly for the subgroup of older adults (aged  $\geq 65$  years). Residence in higher-SES neighborhoods (Q2, Q3, and Q4) was associated with a lower risk of becoming a high-cost utilizer in both the age-gender-SDOH model and the claims history-SDOH model. The results for the other 2 age groups were mixed, with less precise estimates.

## DISCUSSION

This study examined whether SDOH can add incremental value in predicting high-cost utilizers among a commercially insured population from a large national insurer. Overall, the SDOH added minimal predictive power to models that include either just basic demographic characteristics (age and gender) or full

**FIGURE 2.** Ten-Fold ROC by Age-Gender, Age-Gender-SDOH, Full Claims, and Full Claims-SDOH Models for the Full Analysis Sample



AUC, area under the curve; ROC, receiver operating characteristic; SDOH, social determinants of health.



claims information. However, our regression results show that a composite index of SDOH was significantly associated with becoming a high-cost utilizer. We found that residence in a higher-SES neighborhood is associated with a lower risk of becoming a high-cost utilizer for the overall sample population and the population of older adults. However, we found that among children and younger adults (aged 19-44 years), residence in a higher-SES neighborhood is associated with an increased risk of becoming a high-cost utilizer.

To our knowledge, this study is the first to evaluate the relationship between residence in more disadvantaged neighborhoods and the risk of becoming a high-cost utilizer across various age groups within the more homogenous population of commercially insured individuals. Most other studies lack the ability to perform subsample analysis due to the scope of data available in administrative databases (ie, only Medicare or only Medicaid). Our study demonstrates that even though the addition of SDOH did not improve the prediction of high-cost status, poor neighborhood SES and quality of housing/environmental conditions are associated with higher risk of becoming a high-cost utilizer, especially among older adults. The adverse effect of higher SES on the probability of becoming a high-cost utilizer among the full sample and the subsample of older adults is consistent with the existing literature.<sup>12</sup> Individuals from lower-SES groups tend to be sicker, and they use health care services more frequently, thus becoming more high cost.<sup>23</sup>

The major insight of our study is that residence in higher-SES neighborhoods is positively associated with the risk of high-cost utilization among children and younger adults, although the effect is less precisely estimated in the full model for the populations aged 19 to 44 years and 45 to 64 years. Individuals younger than 44 years residing in the highest-ranked neighborhoods by SES are at a higher risk of becoming high-cost utilizers, whereas older individuals residing in higher-SES neighborhoods are at a lower risk. Commercially insured children and younger adults residing in higher-SES neighborhoods might have better access to health care services than their more disadvantaged counterparts. This relative ease of access may

**TABLE 2.** Association Between SES Category and High-Cost Utilizer Status: Full Sample<sup>a</sup>

Variables	Full sample			
	Age-gender-SDOH model		Claims history-SDOH model	
	Percentage-point difference (robust SE)	P	Percentage-point difference (robust SE)	P
	N = 3,101,259		N = 3,101,259	
Age	0.244 (0.001)	<.001	0.029 (0.001)	<.001
Female	2.354 (0.033)	<.001	0.648 (0.032)	<.001
SES category Q2	-0.064 (0.058)	.273	-0.094 (0.055)	.087
SES category Q3	-0.502 (0.057)	<.001	-0.328 (0.053)	<.001
SES category Q4	-0.677 (0.055)	<.001	-0.577 (0.052)	<.001
Specialist visit, n			1.161 (0.061)	<.001
Routine physical visit, n			-2.219 (0.044)	<.001
ED visit, n			2.430 (0.134)	<.001
Inpatient stay, n			-1.231 (0.204)	<.001
Office visit, n			1.673 (0.051)	<.001
Other outpatient visit, n			1.108 (0.049)	<.001
Rx fill, n			-1.662 (0.031)	<.001
Wellness visit, n			-19.699 (0.334)	<.001
Log total specialist cost			0.111 (0.027)	<.001
Log total inpatient stay cost			0.316 (0.018)	<.001
Log total ED cost			0.074 (0.016)	<.001
Log total other outpatient cost			0.738 (0.021)	<.001
Log total office visit cost			-0.637 (0.021)	<.001
Log total wellness cost			2.345 (0.052)	<.001
Log total Rx cost			1.739 (0.012)	<.001
Modified Charlson Comorbidity Index score			0.831 (0.014)	<.001
Alzheimer disease			-3.467 (0.220)	<.001
Arthritis			2.396 (0.064)	<.001
Asthma			0.208 (0.077)	.007
Congestive heart failure			0.327 (0.125)	.009
Chronic obstructive pulmonary disease			1.338 (0.105)	<.001
Cancer			0.517 (0.078)	<.001
Cardiovascular			-0.717 (0.066)	<.001
Crohn disease			21.323 (0.477)	<.001
Depression			1.076 (0.062)	<.001
Diabetes			3.577 (0.066)	<.001
Adrenal insufficiency			-0.385 (0.098)	<.001
Epilepsy			5.948 (0.242)	<.001
Hepatitis C virus			2.420 (0.295)	<.001
HIV			37.835 (0.828)	<.001
Multiple sclerosis			28.314 (0.582)	<.001
Mental illness			3.722 (0.343)	<.001
Pregnancy			-4.405 (0.109)	<.001
Schizophrenia			3.142 (0.560)	<.001
Sickle cell disease			7.519 (1.209)	<.001
Stroke			-1.211 (0.180)	<.001

ED, emergency department; Q, quartile; Rx, prescription; SDOH, social determinants of health; SES, socioeconomic status.

<sup>a</sup>Each column represents a probit regression. The dependent variable equals 1 if an individual is a top 10% high-cost utilizer, and 0 otherwise. All columns include state fixed effects. The results are relative to Q1 (most disadvantaged neighborhood). All coefficients are marginal effects.

**TABLE 3.** Association Between SES Category and High-Cost Utilizer Status: Subgroup Analysis by Age\*

Variables	Age-gender-SDOH model		Claims history-SDOH model	
	Marginal effect: percentage point difference (robust SE)	P	Marginal effect: percentage point difference (robust SE)	P
<b>Children (aged 0-18 years)</b>				
Observations	n = 610,960		n = 610,959	
SES category Q2	0.588 [0.157]	<.001	0.165 [0.144]	.253
SES category Q3	0.898 [0.151]	<.001	0.089 [0.138]	.520
SES category Q4	2.276 [0.144]	<.001	0.548 [0.133]	<.001
<b>Adults (aged 19-44 years)</b>				
Observations	n = 1,127,322		n = 1,127,320	
SES category Q2	0.331 [0.098]	<.001	0.12 [0.09]	.198
SES category Q3	0.067 [0.095]	.484	-0.01 [0.09]	.923
SES category Q4	0.254 [0.092]	<.001	-0.18 [0.09]	.039
<b>Adults (aged 45-64 years)</b>				
Observations	n = 1,042,200		n = 1,042,200	
SES category Q2	-0.248 [0.100]	.013	0.197 [0.092]	.031
SES category Q3	-0.861 [0.098]	<.001	0.110 [0.090]	.221
SES category Q4	-1.477 [0.095]	<.001	0.067 [0.089]	.451
<b>Adults (aged 65-89 years)</b>				
Observations	n = 320,777		n = 320,777	
SES category Q2	-0.709 [0.185]	<.001	-0.487 [0.175]	.005
SES category Q3	-1.341 [0.178]	<.001	-0.744 [0.169]	<.001
SES category Q4	-1.792 [0.174]	<.001	-0.898 [0.167]	<.001

Q, quartile; SDOH, social determinants of health; SES, socioeconomic status.  
 \*Each column represents a probit regression. The dependent variable equals 1 if an individual is a top 10% high-cost utilizer, and 0 otherwise. All columns include state fixed effects. The results are relative to Q1 (most disadvantaged neighborhood). All coefficients are marginal effects.

lead to accumulation of higher health care costs. On the other hand, older adults residing in more disadvantaged communities might be unobservably sicker, leading to higher utilization and higher costs. In other words, better access to health care is likely driving the probability of becoming a high-cost utilizer in younger people, whereas burden of disease is likely driving it for the older adults. These results suggest that the mechanisms behind the relationship between SES and high-cost utilizer status may differ among the age groups.

Our findings help inform a number of payer policies, especially regarding the Medicare Advantage population, whose health care utilization appears to be most correlated with their SDOH. As enrollment in Medicare Advantage plans continues to increase, it is crucial for insurance companies to recognize the importance of SDOH in older individuals' health status and subsequent health care utilization and costs. Medicare Advantage plans may offer supplemental services to address unmet social needs, such as healthy foods and transportation, in addition to medical care to manage members' health, especially in more disadvantaged neighborhoods

with lower SES. Our findings also support the growing trend among health insurers to utilize a more holistic approach in assessing individuals' well-being, focusing on their overall wellness and not just their health condition.

### Limitations

This study has several limitations. First, the analysis included only individuals from a commercially insured population, who on average reside in less disadvantaged communities. As such, our results do not have much external validity for Medicaid or uninsured populations. Other studies have been able to look at either only Medicaid or Medicare, which makes our study more relevant for the US adult working population and their dependents. Second, the variables estimating SDOH represent 5-year estimates (2012-2017), as more precise annual estimates are not publicly available at the desired geographic level (Census block group). Third, because our sample excludes cancer costs, the results may not be generalizable to patients with cancer. Future studies should concentrate on identifying the relationship between SDOH and probability of being a high-cost utilizer among individuals with a specific condition (eg, cancer, diabetes). Fourth, we do not have data on individual-level race/ethnicity, which might help explain disparities in high-cost utilizer status.

## CONCLUSIONS

As health care costs continue to grow, numerous efforts are underway in health care organizations across the nation to identify the individuals at risk of becoming high-cost health care utilizers. There is a recognition among providers, insurers, policy makers, and other stakeholders that socioeconomic background and quality of housing/environmental conditions may be correlated with the need for high-priced health care services, especially among older adults. Our study demonstrates that SDOH are in fact significantly associated with the likelihood of becoming a high-cost utilizer among commercially insured and Medicare Advantage individuals. Localized as well as national policies that recognize the importance of social upbringing and living situation in one's health outcomes are urgently needed. ■

**Author Affiliations:** School of Public Health, Texas A&M University (EA), College Station, TX; Elevance Health (WC), Wilmington, DE; Department of Population Health Sciences, Weill Cornell Medical College (YZ), New York, NY; Weill Cornell Medicine and NewYork-Presbyterian Hospital (RK), New York, NY; Janssen (KH), Wayne, PA.

**Source of Funding:** This study was funded through a Patient-Centered Outcome Research Institute Award (HP-1510-32545) for development of the National Patient-Centered Clinical Research Network, known as PCORnet.

**Author Disclosures:** Dr Chi reports that this work was carried out as part of her employment at Elevance Health, Inc. Drs Zhang and Kaushal are inventors of a pending provisional patent about categorizing and predicting patients with high preventable health care cost. The remaining authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

**Authorship Information:** Concept and design (EA, WC, YZ, RK, KH); acquisition of data (EA, WC, KH); analysis and interpretation of data (EA, WC, KH); drafting of the manuscript (EA, WC); critical revision of the manuscript for important



intellectual content (WC, YZ, RK, KH); statistical analysis (EA, WC); obtaining funding (KH); administrative, technical, or logistic support (WC, YZ, RK); and supervision (YZ, RK).

**Address Correspondence to:** Elena Andreyeva, PhD, Department of Health Policy and Management, School of Public Health, Texas A&M University, 212 Adriance Lab Rd, 1266 TAMU, College Station, TX 77843-1266. Email: andreyeva@tamu.edu.

## REFERENCES

- Agarwal R, Gupta A, Fendrick AM. Value-based insurance design improves medication adherence without an increase in total health care spending. *Health Aff (Millwood)*. 2018;37(7):1057-1064. doi:10.1377/hlthaff.2017.1633
- Casalino LP, Khullar D. Value-based purchasing and physician professionalism. *JAMA*. 2019;322(17):1647-1648. doi:10.1001/jama.2019.14990
- Figuroa JF, Horneffer KE, Jha AK. Disappointment in the value-based era: time for a fresh approach? *JAMA*. 2019;322(17):1649-1650. doi:10.1001/jama.2019.15918
- Hay JW. Now is the time for transparency in value-based healthcare decision modeling. *Value Health*. 2019;22(5):564-569. doi:10.1016/j.jval.2019.01.009
- Ma Q, Sylwestrzak G, Oza M, Garneau L, DeVries AR. Evaluation of value-based insurance design for primary care. *Am J Manag Care*. 2019;25(5):221-227.
- Perez SL, Gosdin M, Pintor JK, Romano PS. Consumers' perceptions and choices related to three value-based insurance design approaches. *Health Aff (Millwood)*. 2019;38(3):456-463. doi:10.1377/hlthaff.2018.05048
- Pampel FC, Krueger PM, Denney JT. Socioeconomic disparities in health behaviors. *Annu Rev Sociol*. 2010;36:349-370. doi:10.1146/annurev.soc.012809.102529
- Strickhouser JE, Sutin AR. Family and neighborhood socioeconomic status and temperament development from childhood to adolescence. *J Pers*. 2020;88(3):515-529. doi:10.1111/jopy.12507
- Tamura K, Langerman SD, Ceasar JN, Andrews MR, Agrawal M, Powell-Wiley TM. Neighborhood social environment and cardiovascular disease risk. *Curr Cardiovasc Risk Rep*. 2019;13(4):7. doi:10.1007/s12170-019-0601-5
- Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. The National Academies Press; 2003.
- Institute of Medicine. *The Future of the Public's Health in the 21st Century*. The National Academies Press; 2002.
- Fitzpatrick T, Rosella LC, Calzavara A, et al. Looking beyond income and education: socioeconomic status gradients among future high-cost users of health care. *Am J Prev Med*. 2015;49(2):161-171. doi:10.1016/j.amepre.2015.02.018
- Daniel H, Bornstein SS, Kane GC, et al; Health and Public Policy Committee of the American College of Physicians. Addressing social determinants to improve patient care and promote health equity: an American College of Physicians position paper. *Ann Intern Med*. 2018;168(8):577-578. doi:10.7326/M17-2441
- Hill J, Nielsen M, Fox MH. Understanding the social factors that contribute to diabetes: a means to informing health care and social policies for the chronically ill. *Perm J*. 2013;17(2):67-72. doi:10.7812/TPP/12-099
- Sheets L, Popejoy L, Aprn GB, Khalilia M, Petroski G, Parker JC. Identifying patients at risk of high healthcare utilization. *AMIA Annu Symp Proc*. 2017;2016:1129-1138.
- Adler NE, Cutler DM, Fielding JE, et al. Addressing social determinants of health and health disparities. National Academy of Medicine. September 19, 2016. Accessed August 20, 2019. <https://nam.edu/wp-content/uploads/2016/09/Addressing-Social-Determinants-of-Health-and-Health-Disparities.pdf>
- Social determinants of health. HealthyPeople.gov. Accessed August 20, 2019. <https://wayback.archive-it.org/5774/2020413203948/https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health>
- Wodchis WP, Austin PC, Henry DA. A 3-year study of high-cost users of health care. *CMAJ*. 2016;188(3):182-188. doi:10.1503/cmaj.150064
- Murphy JF, Hepworth JT. Age and gender differences in health services utilization. *Res Nurs Health*. 1996;19(4):323-329. doi:10.1002/(SICI)1098-240X(199608)19:4<323::AID-NUR6>3.0.CO;2-K
- Berkman LF, Kawachi I, Glymour MM. *Social Epidemiology*. 2nd ed. Oxford University Press; 2014.
- Bonito AJ, Bann C, Eicheldinger C, Carpenter L. *Creation of New Race-Ethnicity Codes and Socioeconomic Status (SES) Indicators for Medicare Beneficiaries: Final Report*. Agency for Healthcare Research and Quality; 2008.
- Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011;173(6):676-682. doi:10.1093/aje/kwq433
- Droomers M, Westert GP. Do lower socioeconomic groups use more health services, because they suffer from more illnesses? *Eur J Public Health*. 2004;14(3):311-313. doi:10.1093/eurpub/14.3.311

Visit [ajmc.com/link/89399](https://ajmc.com/link/89399) to download PDF